

Building a Hypertextual Digital Library in the Humanities: A Case Study on London

*Gregory Crane, David A. Smith,
Clifford E. Wulfman*

Perseus Project
Eaton Hall
Tufts University
Medford MA 02155

“Humanities” DL Issues

- The DL *is* the laboratory
 - Citation PLUS page # standard
 - Very fine grained citation schemes
 - Documents rule, style sheets serve
 - Descriptive rather than prescriptive

Maturity of Print Matter

- Little innovation in Print Form since 1900
- Advantage: PD Materials useful as testbed
- Disadvantage: Little work on the *form* of documents
 - Inadequate knowledge base for innovation
 - Worse: little place for thinking about such problems in the academy
 - (Humanist Critical traditions weak in this area.)

Designing a Hum DL

- The Critic vs. the Editor
 - Production of scholarly infrastructure complex
 - Editions/Commentaries/Ref Works Crucial
- The Problem:
 - How to design long term docs until we know what DLs can and should do?
- DL challenges us to rethink:
 - Documents and what they can do
 - Our Audience: Other researchers? General Public?

Depth First Approach

- Perseus Greco-Roman Collections
 - Major work begun in 1987
 - Portability: CD ROM 1992, WWW 1996
 - Language Analysis, Complex Citation Mining
 - “Hard” problems but obvious wins
 - Sparsity of Data: “Desert agriculture”
 - Popular: 67m pages, 6.7m sessions in 2000



Classics:

[Collection contents](#)
[About the collection](#)

► [Greek Hist. Overview](#)
 ► [Art & Arch. Catalogs](#)

Latin Tools:

► [Grammar Overview](#)
 ► [Dictionaries](#)
 ► [Morphology](#)
 ► [Word Search](#)

► [Other Tools & Lexica](#)

Contents:

- [Book 1: LIBER I](#)
- [Book 2: LIBER II](#)
- [Book 3: LIBER III](#)
- [Book 4: LIBER IV](#)
- [Book 5: LIBER V](#)

Collections: [Classics](#) ? [Papyri](#) ? [Renaissance](#) ? [London](#) ? [California](#) ? [Upper Midwest](#) ? [Tufts History](#)

[Configure display](#) ? [Help](#) ? [Tools](#) ? [Copyright](#) ? [FAQ](#) ? [Publications](#) ? [Collaborations](#) ? [Support Perseus](#)

Cornelius Tacitus, *Historiae*

LIBER II

Version:

Your current position in the text is marked in

← Go to

CI. Scriptores temporum, qui potente
 publicae, corruptas in adulationem caus
 etiam invidiaque, ne ab aliis apud Vitell
 centurionum militumque animos obstin
 lubrica ad mutandam fidem classe ob n

Preferred URL for linking to this page: <http://www.perseus.tufts.edu/cgi-bin/ptext?lookup=Tac.+Hist.+2.101>

This text is based on the following book(s):

Historiae. Cornelius Tacitus. Charles Dennis Fisher. Clarendon Press. Oxford. 1911.

OCLC: 13976624

Netscape: Word Study Tool

corrumpto

to destroy, ruin, waste

Entry in [L&S](#)

corrupt?s

perf part pass fem acc pl

Frequency by Authors

Latin Word Search

Corpus	Words	Max. Inst.	Freq./10K	Min. Inst.	Freq./10K
Latin Texts	1792745	221	1.23	101	0.56

corruptus

spoiled, marred, corrupted, bad

Entry in [L&S](#)

corrupt?s

fem acc pl

Frequency by Authors

Latin Word Search

Corpus	Words	Max. Inst.	Freq./10K	Min. Inst.	Freq./10K
Latin Texts	1792745	120	0.67	0	0



Breadth Second Study

- DLI2: Case Studies in Humanities DLs
 - Ancient Egypt with Boston MFA
 - History of Mechanics: MPI Berlin (NSF/DFG)
 - History of Recent Science: MIT Dibner
 - New Variorum Shakespeare Series: MLA
 - American Memory Collections: LOC
- Domain Specific vs. General Problems

A DL on London: Preliminary Results

- Edwin Bolles Collection at Tufts
- Intensive Discourse / Limited Spatial Focus
- Centuries of Data
- Largely Monolingual
- Books, Images, Historic Maps
 - Paper Hypertext, Geo-referencing of maps

ProtoWeb Browser

Note: Blue “Links”

Monument of Capt. Rundle Burgess.

St. Paul's.]

NELSON'S FUNERAL IN ST. PAUL'S.

251

1789, George III. came to thank God for his temporary recovery from insanity. Queen Charlotte, the Prince of Wales, and the Duke of York were present, and both Houses of Parliament. Bishop Porteous preached the sermon, and 6,000 charity children joined in the service. In 1797, King George came again to attend a thanksgiving for Lord Duncan's and Lord Howe's naval victories; French, Spanish, and Dutch flags waved above the procession, and Sir Horatio Nelson was there among other heroes.

The first grave sunk in St. Paul's was fittingly that of Wren, its builder. He lies in the place of honour, the extreme east of the crypt. The black marble slab is railed in, and the light from a small window-grating falls upon the venerated name. Sir Christopher died in 1723, aged ninety-one. The fine inscription, "Si monumentum requiris, circumspice," written probably by his son, or Mylne, the builder of Blackfriars Bridge, was formerly in front of the organ-gallery, but is now placed over the north-western entrance.

The clergy of St. Paul's were for a long time jealous of allowing any monument in the cathedral. Dean Newton wished for a tomb, but it was afterwards erected in St. Mary-le-Bow. A better man than the vain, place-hunting dean was the first honoured. The earliest statue admitted was that of the benevolent Howard, who had mitigated suffering and sorrow in all the prisons of Europe; he stands at the corner of the dome facing that half-stripped athlete, Dr. Johnson, and the two are generally taken by country visitors for St. Peter and St. Paul. He who with Goldsmith had wandered through the Abbey, wondering if one day their names might not be recorded there, found a grave in Westminster, and, thanks to Reynolds, the first place of honour. Sir Joshua himself, as one of our greatest painters, took the third place, that Hogarth should have occupied; and the fourth was awarded to that great Oriental scholar, Sir William Jones. The clerical opposition was now broken through, for the world felt that the Abbey was full enough, and that St. Paul's required adorning.

Henceforward St. Paul's was chiefly set apart for naval and military heroes whom the city could best appreciate, while the poets, great writers, and statesmen were honoured in the Abbey, and laid among the old historic dead. From the beginning our sculptors resorted to pagan emblems and pagan allegorical figures; the result is that St. Paul's resembles a Pantheon of the Lower Empire, and is a hospital of third-rate art. The first naval conqueror so honoured was Rodney; Rossi received £6,000 for his cold and clumsy design:

Lord Howe's statue followed; and next that of Lord Duncan, the hero of Camperdown. It is a simple statue by Westmacott, with a seaman and his wife and child on the pedestal. For Earl St. Vincent, Bailey produced a colossal statue and the usual scribbling, History and a trumpeting Victory.

Then came Nelson's brothers in arms—men of lesser mark; but the nation was grateful, and the Government was anxious to justify its wars by its victories. St. Paul's was growing less particular, and now opened its arms to the best men it could get. Many of Nelson's captains preceded him on the red road to death—Westcott, who fell at Aboukir; Mosse and Riou, who fell before Copenhagen (a far from stainless victory). Riou was the brave man whom Campbell immortalised in his fiery "Battle of the Baltic." Riou lies

"Full many a fathom deep,
By thy wild and stormy sleep,
Elsinore."

Then at last, in 1806, came a hero worthy, indeed, of such a cathedral—Nelson himself. At what a moment had Nelson expired! At the close of a victory that had annihilated the fleets of France and Spain, and secured to Britain the empire of the seas. The whole nation that day shed tears of "pride and of sorrow." The Prince of Wales and all his brothers led the procession of nearly 8,000 soldiers, and the chief mourner was Admiral Parker (the Mutiny of the *Nore* Parker). Nelson's coffin was formed out of a mast of the *L'Orient*—a vessel blown up at the battle of the Nile, and presented to Nelson by his friend, the captain of the *Swiftsure*. The sarcophagus, singularly enough, had been designed by Michael Angelo's contemporary, Torreguano, for Wolsey, in the days of his most insatiable pride, and had remained ever since in Wolsey's chapel at Windsor; Nelson's flag was to have been placed over the coffin, but as it was about to be lowered, the sailors who had borne it, as if by an irresistible impulse, stepped forward and tore it in pieces, for relics. Dean Milman, who, as a youth, was present, says, "I heard, or fancied I heard, the low wail of the sailors who encircled the remains of their admiral." Nelson's trusty companion, Lord Collingwood, who led the vanguard at Trafalgar, sleeps near his old captain, and Lord Northesk, who led the rear-guard, is buried opposite. A brass plate on the pavement under the dome marks the spot of Nelson's tomb. The monument to Nelson, inconveniently placed at the opening of the choir, is by one of our greatest sculptors—Flaxman. It is hardly worthy of the occasion, and the figures on the pedestal are puerile. Lord

10 Richard Rundle Burgess

Scholarly Reading

- Reading as Information Extraction
 - Decoding and absorbing well-defined msgs
- Reading as Literary Play
 - Postmodern theory: popular in HT work
- The historicist scholar
 - Partakes of **both** extremes at once
 - Active tension rather than simple balance

Scholarly Reading

- Documents are NOT transparent
 - We read *against* the grain of the documents
 - E.g., Civil War songs as sources on Gender
 - Even historians are themselves objects of study
 - Constant shifting between fast/intensive reading
 - Contextualization is crucial
 - Look up the footnotes, read the sources, find new sources

Consequences for HUM DLs: 1

- DLs not aimed at short-term return
 - Persistence of Reference paramount
 - Corporate repositories problematic
- Design for widest possible audience
 - Good structure rather than oversimplification
- Size Matters: We need LOTS of data
- Need fine granularity of reference

Consequences for HUM DLs: 2

- We need dense, flexible networks of links
 - We need *automatic* linking technologies
 - We need the ability to *refine* auto-links
- Critical Question in the Humanities:
 - How far can the machines bring us?
 - Where does scarce labor begin?
 - How do the machines change the goals of labor
- E.g., Do we need a new “narrative” dictionary or semantic browser?

Links in the London DL

- Latin (and Greek) Sections
 - DL automatically applies language tools
 - Searching, indexing, lexicon links
 - Benefit of a single extensible environment
- Named Entity Id#: People/Places/Money etc
 - Plenty of research on this problem
 - But how to bootstrap authority lists?
 - Yes, “Oliver Cromwell” is a person, but which one?
 - Semantic vs. Encyclopedic Reference!

Bootstrapping Authority Lists

- Cruchley 1843; Bartholemew 1999
 - 48,000 of 90,000 (53%) Streets in Bolles in Barth.
- Dictionary of National Biography Summary
- Wheatley's Encyclopedia of London
- Various guides and surveys:
 - Headwords of Chapters and Sections Mined
- Image Captions

Mapping Various Structures

- Headwords: DNB etc.
 - “Smith, Sir John Evelyn” etc.
- Chapter and Section Headings:
 - “Chapter 2: Tower of London”
- Embedded Keywords
 - “Now we turn to the *Tower of London* as we..”
 - Italics used to mark subject strings

Applying Authority Links

- Simple Pattern Matching to Add Links
 - Assumes coherence of patterns and docs
 - Naïve application surprisingly effective

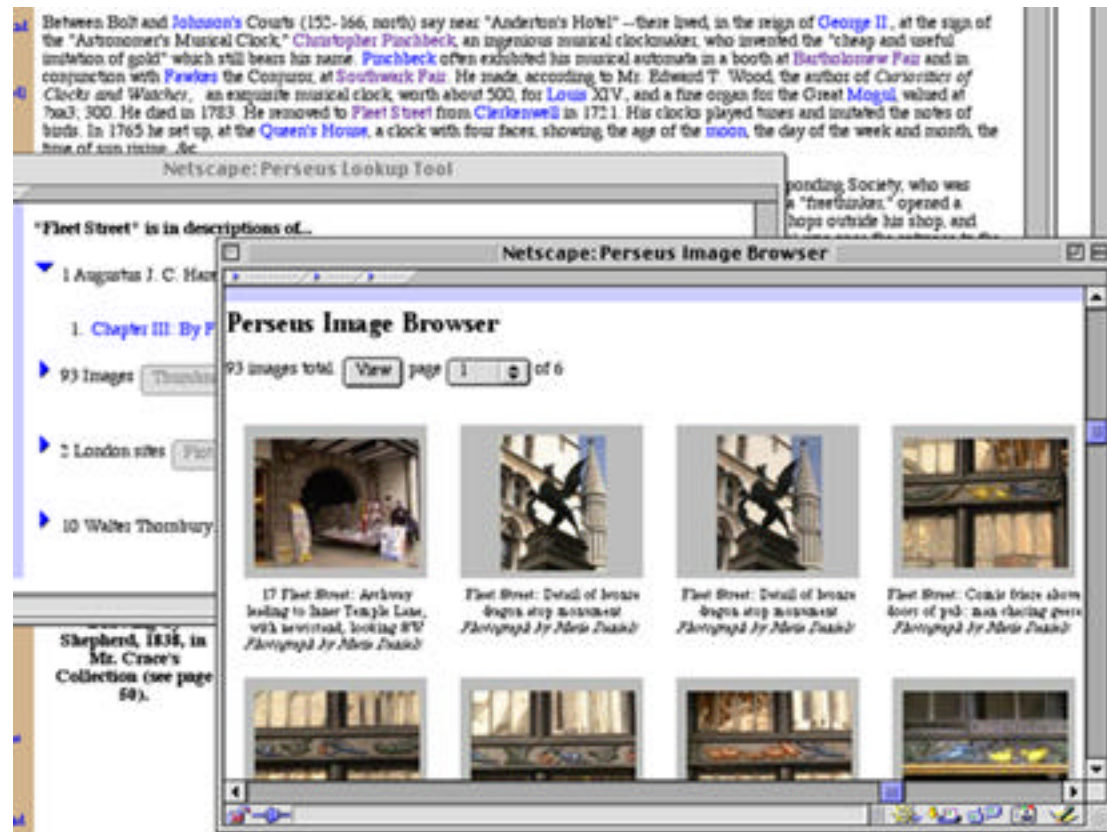
Bottom Up 1: Automatic Links

- There were certainly rough doings in Fleet Street in the Middle Ages, for the City chronicles tell us of much blood spilt there and of many deeds of violence. In 1228 (Henry III.) we find, for instance, one Henry de Buke slaying a man named Le Ireis, or Le Tylor, of Fleet Bridge, then fleeing to the church of St. Mary, Southwark, and there claiming sanctuary. In 1311 (Edward II.) five of the king's not very respectable or law-fearing household were arrested in Fleet Street for a burglary;

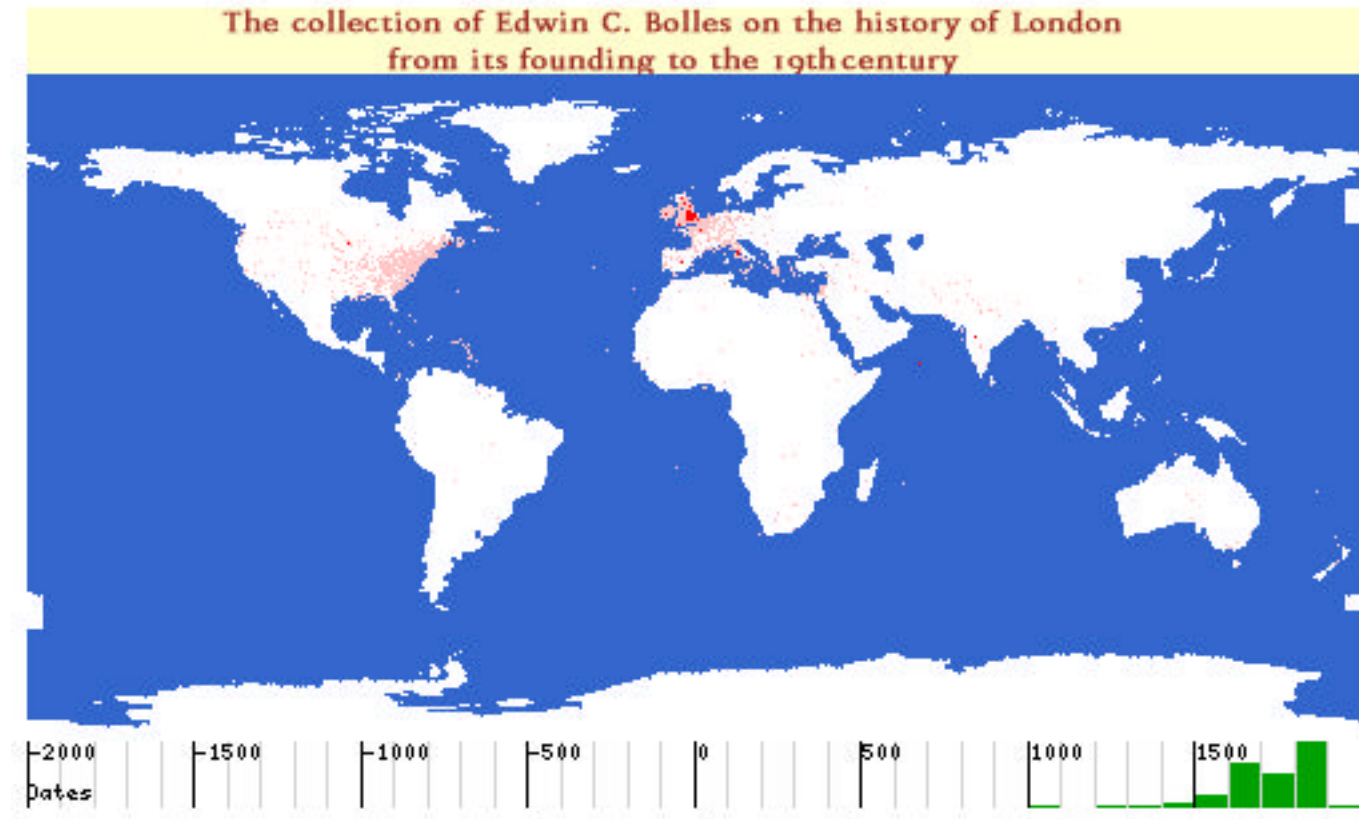
Bottom Up 2: Aggregation

Fleet Street'' is in descriptions of...
1 Hare Chapter
By Fleet Street to St. Paul's., Fleet Street
95 Images
2 London sites
1.Fleet Street
2.Fleet Street Hill
10 Thornbury chapters

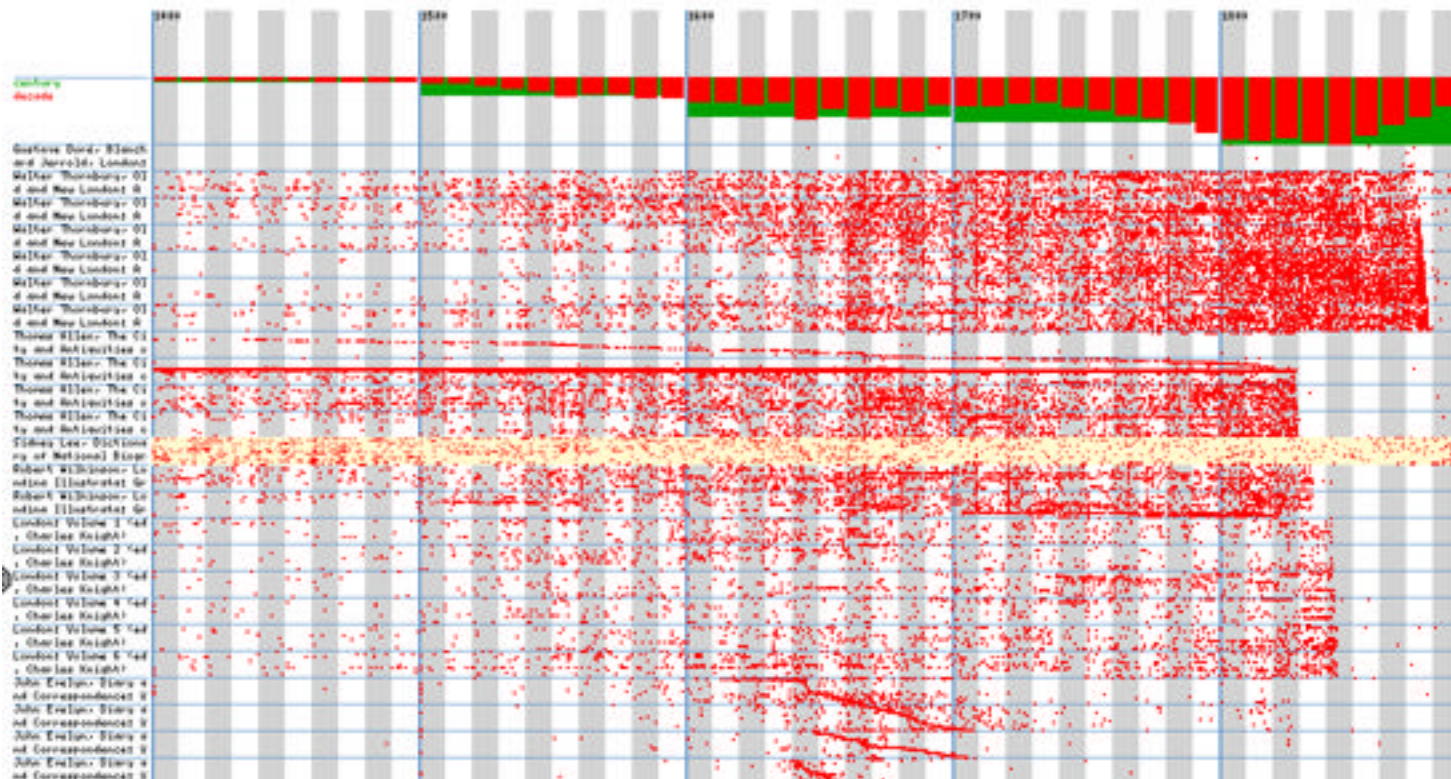
Bottom Up 3: Aggregation



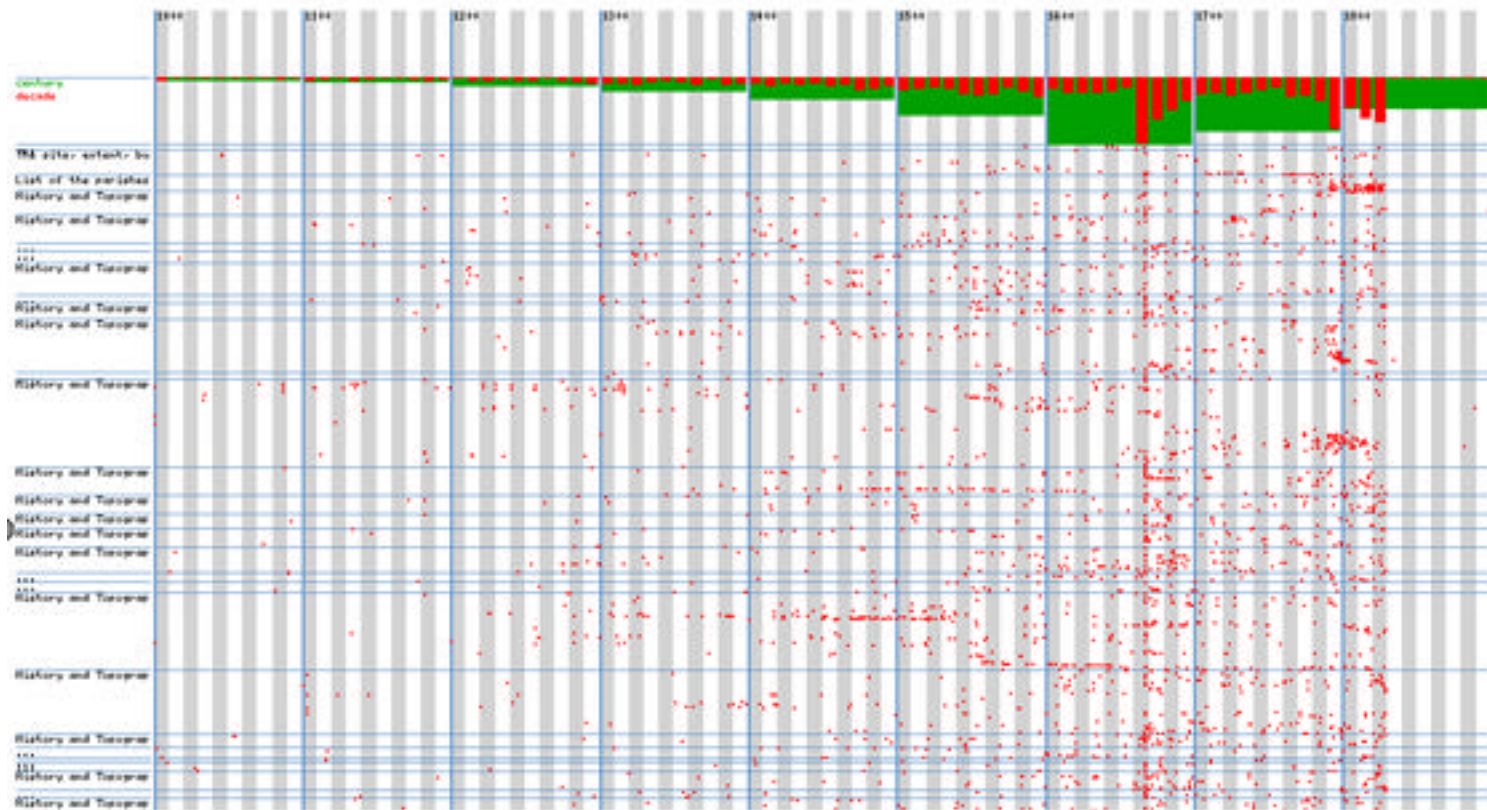
Top Down 1: Time & Space



Top Down 2: Automatic Timeline



Top Down 3: Book Timeline



Top Down 4: Mapping

Walter Thornbury, *Old and New London: A Narrative of its History, its People and its Places. Illustrated with Numerous Engravings from the Most Authentic Sources*

Your current position in the text is marked by a red line.

Go to: 54

Plot streets on this page
(p. 54)

Fleet Street (continued).

- The "Green Dragon"
- The "Boh-in-Tun"
- [Tongson and Patchbeck](#)
- The *Record*
- St. Bride's and its Menories
- Patch and his Contributions
- The *Daily Mail*
- The *Daily Telegraph*
- The "Globe Tavern" and Old
- The *Morning Advertiser*
- The *Standard*
- The *London Magazine*
- A Strange Story
- [Alderman Woburn](#)
- [Bryon Billy](#)
- [Hardham](#) and his "37"

The original "Green Dragon" (No. 56) see the shouting, tooth-aching process when young men.

No. 64 (north) was [long](#), a well-known years ago. Its name is preserved in Bo.

At No. 67 (corner of [Whitehall](#) Street) without winding up. He died in 1713 with. The old shop was (1850) one of

Between Bolt and Johnson's Courts (1 clockmaker, who invented the "cheap 4 [Southwick Fox](#). He made, according to had 3, 300. He died in 1783. He removed the [moon](#), the day of the week and mo

No. 161 (north) was the shop of [Thos](#) Richard [Culark](#), a "feartholder," open imprisonment a punishment by no means of his admirers, to whom he re

"Anderson's Hotel" (No. 164, north) registers of St. Dunstons is an entry (1 against [Sergeant](#) [Lau](#), and Thos-jeggs

The *Record* (No. 169, north) was started in 1816 as an organ of the extreme Evangelical party. The first promoters were the late Mr. James [Evans](#), a brother of Sir [Andrew Agnew](#), and Mr. And

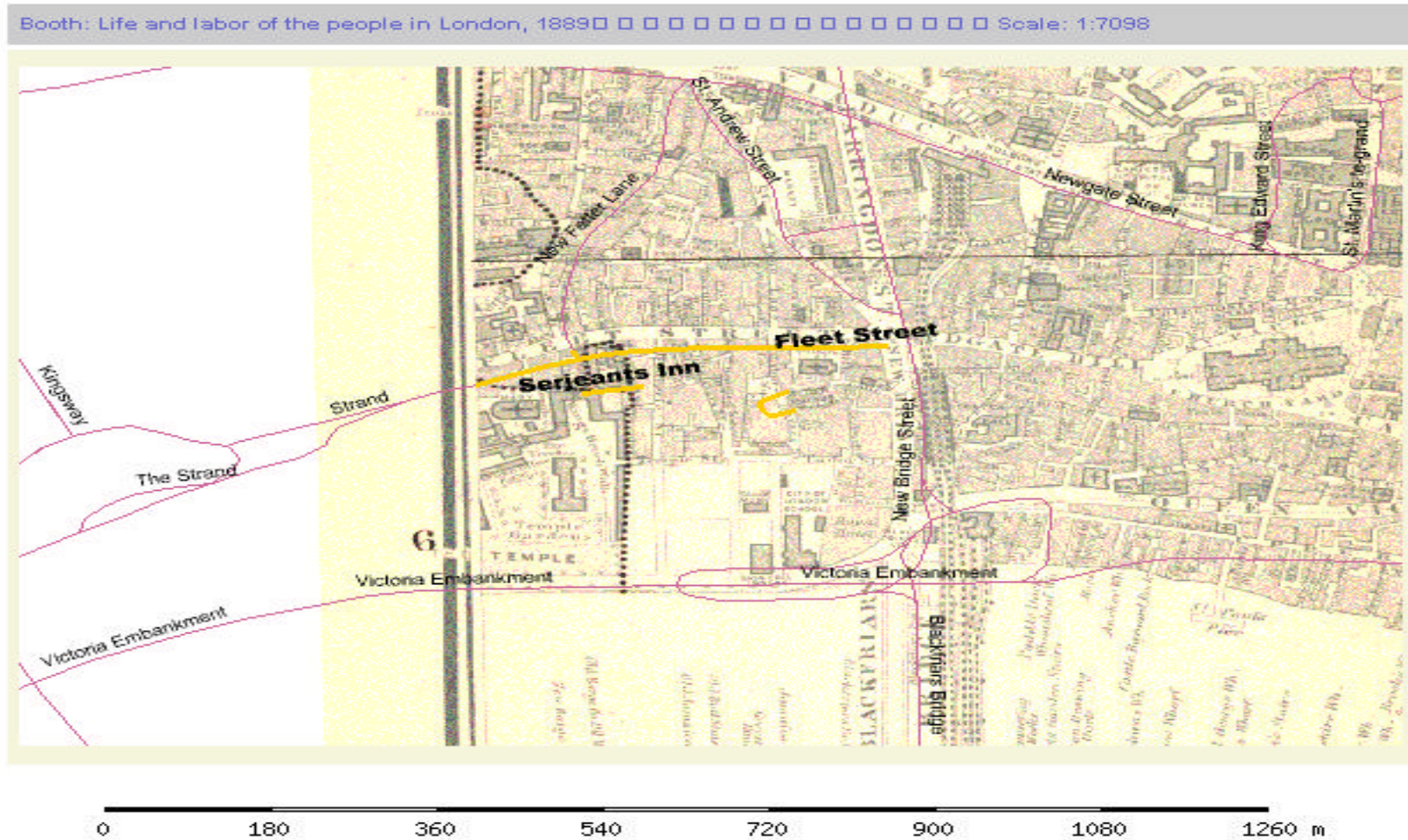
Netscape: Balfes Collection London Map Browser

Location: <http://www.person.hkft.edu/cgi-bin/city-view.pl?scale=14136.89749502645&addstreet=fleet+stre>

City View of Modern London

The map shows a section of London, including Fleet Street, Serpentine Inn, and various other landmarks. The map is titled "City View of Modern London". The browser window has a menu bar with options like Back, Forward, Reload, Home, Search, Netscape, Images, Print, Security, Shop, and Stop. The address bar shows the URL: <http://www.person.hkft.edu/cgi-bin/city-view.pl?scale=14136.89749502645&addstreet=fleet+stre>. The browser window also has a search bar and a list of links.

Geospatial Links: GIS to Map



Future Work

- Other sources of link data: Arguments and Conventional Indices
- Quotes and Citation Linking
- Tabular Information
- Monetary sums
- Temporal Spatial Querying
- Providing Link Services to External Datasets

Conclusions

- Bootstrapping a Hum DL:
 - Start with a knowledge base
 - Convert Ref Works and Structured info first
 - Dictionaries/encyclopedias/etc. first
- Integration of Knowledge Sources Essential
 - Automatic disambiguation
 - Editing Environment needed
- Determining the Form of Cultural Heritage Information is key to any comprehensive DL